DOCUMENT RESUME

ED 079 352 TM 002 951

AUTHOR	Stroud, T. W. F.
TITLE	Comparing Regressions When Measurement Error Variances Are Known.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RB-73-35
PUB DATE	May 73
NOTE	31p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Achievement Tests; *Comparative Analysis; Hypothesis Testing; *Mathematical Models; *Multiple Regression Analysis; Research Methodology; Secondary Grades; Standard Error of Measurement; Technical Reports; *Test Results; *True Scores

ABSTRACT

In a multiple (or multivariate) regression model where the predictors are subject to errors of measurement with a known variance-covariance structure, two-sample hypotheses are formulated for (i) equality of regressions on true scores and (ii) equality of residual variance (or covariance matrices) after regression on true scores. The hypotheses are tested using a large-sample procedure based on maximum likelihood estimators. Formulas for the test statistic are presented; these may be avoided in practice by using a general purpose computer program. The procedure has been applied to a comparison of learning in high schools using achievement test data. (Author)

ED 079352
TM 002 951

RESEARCH BULLETIN

RB-73-35

COMPARING REGRESSIONS WHEN MEASUREMENT ERROR VARIANCES ARE KNOWN

T. W. F. Stroud
Queen's University
and
Educational Testing Service

This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the author. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
May 1973

FILMED FROM BEST AVAILABLE COPY

COMPARING REGRESSIONS WHEN MEASUREMENT ERROR VARIANCES ARE KNOWN

Abstract

In a multiple (or multivariate) regression model, where the predictors are subject to errors of measurement with a known variance-covariance structure, two-sample hypotheses are formulated for (i) equality of regressions on true scores and (ii) equality of residual variances (or covariance matrices) after regression on true scores. The hypotheses are tested using a large-sample procedure based on maximum likelihood estimators. Formulas for the test statistic are presented; these may be avoided in practice by using a general purpose computer program. The procedure has been applied to a comparison of learning in high schools using achievement test data.

COMPARING REGRESSIONS WHEN MEASUREMENT ERROR VARIANCES ARE KNOWN

1. Introduction

Often we want to compare two groups of subjects on the basis of a posttest, with adjustment made for scores on a pretest. For this purpose many experimenters have used the analysis of covariance. When the observed pretest scores contain errors of measurement, however, we would really like to make our adjustment in terms of true scores [Lord & Novick, 1968], since otherwise the covariance adjustment does not properly correct for the bias in the difference of adjusted means as an estimate of the difference in intercepts [Cochran, 1968, p. 653]. Because true scores are unobservable, one needs extra information to make a satisfactory correction. Lord [1960] has proposed a method of doing this when duplicate measurements (with independent errors) on an individual's score are available.

Another problem frequently encountered is that the within-group regressions may not be parallel. In the case of one pretest variable, the regression lines may cross so that the subjects in one group may score higher or lower on the posttest than those in the other group, depending on the pretest score. Cronbach and Gleser [1965, pp. 177-181] give some examples of this, and Johnson and Jackson [1959, pp. 424ff] describe a method (the Johnson-Neyman technique) for ascertaining statistical significance in this case.

In this article we propose a method of testing for equality of regressions in two groups when all variables in the regression equation are true scores. The end is similar to that achieved by Lord [1960], but we allow

the regression slopes to be different. Instead of assuming that duplicate measurements exist, we assume that the pretest measurement error variance-covariance matrix is known and the same for both groups, and that the posttest measurement error covariance matrix is the same for both groups. Either the pretest score or the posttest score, or both, may be multivariate. The case where both are univariate is described in Stroud [1972].

The study which motivated the methodology described here involved a desire to compare the learning taking place in two groups of schools, where the Iowa Tests of Educational Development (ITED) were administered in grade 9 (pretests) and the Tests of Academic Progress (TAP) were administered in grade 11 (posttests). It was felt this could be best achieved by testing the null hypothesis that the true score regressions were the same for both groups. The formulation described here should be applicable in many situations involving tests or measurements in which multivariate normality may be assumed and measurement error variances and covariances of the predictor variables are known, as is the case with tests such as the ITED if one is willing to use the publisher's figures for standard errors of measurement and assume the subtest measurement errors uncorrelated.

For subjects in the first group, let us write, using the classical test theory model, the vector equation $X = T + E$ (observed score = true score + error of measurement), which we partition as

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} + \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} .$$

Here X_1 is the vector of pretest scores for an individual, X_2 the corresponding vector of posttest scores (not necessarily of the same dimension),

and T_1, T_2, E_1, and E_2 are the corresponding true score and error vectors. We assume zero correlation between T and E and between E_1 and E_2. For subjects in the second group, let the corresponding equation be $Y = U + F$, again partitioned into pretest and posttest vectors.

We are interested in testing the hypothesis that the regression function of posttest true score on pretest true score is the same in both groups, i.e.,

$$H_1: \quad \mathcal{E}[T_2|T_1 = t] = \mathcal{E}[U_2|U_1 = t] \quad \text{for all vectors } t .$$

In the school comparison example cited above, one could interpret this as meaning that the average "learning" (in some sense) is the same in both groups of schools. Another hypothesis of interest is that the learning is equally uniform in both groups of schools, i.e., that the true score residual covariance matrix is the same in both groups:

$$H_2: \quad \mathcal{C}[T_2|T_1 = t] = \mathcal{C}[U_2|U_1 = t] \quad \text{for all vectors } t ,$$

where the symbol $\mathcal{C}$ denotes covariance matrix. We may be interested in either H_1 or H_2 independently of the other; so they are treated separately.

In Section 2 we express hypotheses H_1 and H_2 in terms of the parameters of the observed scores X and Y, assuming multivariate normality and that the measurement errors in the two groups, E and F, have the same covariance matrix with the pretest part known. In Section 3 we

describe how to test these hypotheses using Wald's large-sample test procedure based on maximum likelihood estimators and a computing algorithm [Lord, 1972] which calculates the value of the Wald test statistic without requiring a formula for it, using numerical differentiation. Section 4 contains the results of the above mentioned comparison of schools.

Finally, an appendix is presented which contains a formulation of some properties of the Wald procedure, and formulas for the asymptotic covariance matrix of the maximum likelihood estimator of the left-hand side of each hypothesis, which would enable computation of the Wald statistic without using numerical differentiation.

2. Formulation of Hypotheses in Terms of Parameters of Distribution of Observed Scores

Assume that one is given $m + n$ mutually independent observation vectors; the first m of the form $X = T + E$ (T and E independent) and the remaining n of the form $Y = U + F$, where X and Y are the only quantities observed. Assume all distributions are multivariate normal, viz: $T \sim \mathcal{N}(\mu,\Sigma^*)$, $U \sim \mathcal{N}(\nu,\Psi^*)$, $E \sim \mathcal{N}(0,\Delta)$, and $F \sim \mathcal{N}(0,\Delta)$. X and Y are each partitioned into a pretest part (dimension p) and a posttest part (dimension q), as described in Section 1. This induces a partitioning on μ, ν, Σ^*, Ψ^*, and Δ, e.g.,

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} , \qquad \Sigma^* = \begin{bmatrix} \Sigma^*_{11} & \Sigma^*_{12} \\ \Sigma^*_{21} & \Sigma^*_{22} \end{bmatrix} .$$

ν and Ψ^* are represented similarly. E_1 and E_2 have been assumed uncorrelated; so write $\Delta = \text{diag}(\Delta_1, \Delta_2)$. We assume that Σ^* and Ψ^* are nonsingular, and that Δ_1 is known. In practice Δ_1 will usually be diagonal; however, we allow both Δ_1 and Δ_2 to be any positive semi-definite matrices, provided Δ_1 can be specified.

In regression problems involving errors of measurement, it is common to consider the underlying variables, although random, to be linearly or "structurally" related [Kendall, 1951: Madansky, 1959; Moran, 1971]. We take a different approach and consider rather the unconditional joint distribution of T_1 and T_2 (or of X_1 and X_2). We avoid writing anything as a linear function of T_1; we retain T as $\mathcal{N}(\mu, \Sigma^*)$ throughout and condition on T_1 only to obtain the formulas for $\mathcal{E}[T_2|T_1]$ and $\mathcal{C}[T_2|T_1]$. Taking the formulas for these quantities from Anderson [1958, page 29], H_1 and H_2 may be written as follows:

$$(2.1) \qquad H_1: \begin{cases} \Sigma_{21}^* \Sigma_{11}^{*-1} = \Psi_{21}^* \Psi_{11}^{*-1} \quad , \\ \mu_2 - \Sigma_{21}^* \Sigma_{11}^{*-1} \mu_1 = \nu_2 - \Psi_{21}^* \Psi_{11}^{*-1} \nu_1 \quad , \end{cases}$$

$$(2.2) \qquad H_2: \ \Sigma_{22}^* - \Sigma_{21}^* \Sigma_{11}^{*-1} \Sigma_{12}^* = \Psi_{22}^* - \Psi_{21}^* \Psi_{11}^{*-1} \Psi_{12} \quad .$$

To write these hypotheses in terms of the parameters of the distributions of the observed vectors X and Y, note that X and Y are normally and independently distributed with mean vectors μ, ν and covariance matrices $\Sigma \equiv \Sigma^* + \Delta$ and $\Psi \equiv \Psi^* + \Delta$ respectively. Partition Σ, Ψ in the same way as Σ^*, Ψ^*; then (2.1) and (2.2) become, respectively,

$$(2.3)\qquad H_1:\ \begin{cases} \Sigma_{21}(\Sigma_{11} - \Delta_1)^{-1} = \Psi_{21}(\Psi_{11} - \Delta_1)^{-1} \ , \\ \mu_2 - \Sigma_{21}(\Sigma_{11} - \Delta_1)^{-1}\mu_1 = \nu_2 - \Psi_{21}(\Psi_{11} - \Delta_1)^{-1}\nu_1 \ , \end{cases}$$

$$(2.4)\qquad H_2:\ \Sigma_{22} - \Sigma_{21}(\Sigma_{11} - \Delta_1)^{-1}\Sigma_{12} = \Psi_{22} - \Psi_{21}(\Psi_{11} - \Delta_1)^{-1}\Psi_{12} \ .$$

3. The Hypothesis Testing Procedure

The method proposed in this article for testing H_1 and H_2 is an asymptotic test procedure based on unrestricted maximum likelihood estimators (MLE), first described by Wald [1943]. Consider the general problem of testing a vector hypothesis $g(\mu,\nu,\Sigma,\Psi) = 0$ where (μ,Σ) are the mean vector and covariance matrix of one multivariate normal population and (ν,Ψ) are the mean vector and covariance matrix of a second such population. Denote sample estimates by $(\hat{\mu},\hat{\Sigma})$ and $(\hat{\nu},\hat{\Psi})$ respectively, where the sample sizes are m and n. Let

$$U = g(\hat{\mu},\hat{\nu},\hat{\Sigma},\hat{\Psi}) \ ,$$

written out as a column vector, using each off-diagonal component of the symmetric matrices $\hat{\Sigma}$ and $\hat{\Psi}$ only once. Then the test statistic is

$$(3.1)\qquad W = U'\hat{\Xi}^{-1}U \ ,$$

where $\Xi = \Xi(\mu,\nu,\Sigma,\Psi)$ is a large-sample approximation to the covariance matrix of U, based on the partial derivatives of g, and $\hat{\Xi}$ is defined as $\Xi(\hat{\mu},\hat{\nu},\hat{\Sigma},\hat{\Psi})$.

The estimates $\hat{\mu}$ and $\hat{\nu}$ of the population mean vectors are usually taken as the sample mean vectors $\bar{X}$ and $\bar{Y}$, respectively. $\hat{\Sigma}$ and $\hat{\Psi}$ are the unrestricted maximum-likelihood estimates $\Sigma_{i=1}^{m}(X^i - \bar{X})(X^i - \bar{X})'/m$ and $\Sigma_{i=1}^{n}(Y^i - \bar{Y})(Y^i - \bar{Y})'/n$, although one could instead use the unbiased versions with denominators of $(m - 1)$ and $(n - 1)$, the difference being unimportant in large samples.

The statistic W is asymptotically distributed as central (noncentral) chi-square if the hypothesis is true (false). This result, originally proved by Wald [1943] under rather strict regularity conditions, has been shown by Stroud [1971] to hold under fairly general circumstances, which apply in particular to the problems studied in this article.

To compute the value of W from given data using the computer program described in Lord [1972], it is necessary to write a FORTRAN function to compute each component of U, given $\hat{\mu}$, $\hat{\nu}$, $\hat{\Sigma}$, and $\hat{\Psi}$. For the two problems described here, U is obtained by writing equations (2.3) and (2.4) with estimates substituted in and with all quantities transposed to the left of the equal sign. If H_1 is being tested, the first pq components of U are the components of $\hat{\Sigma}_{21}(\hat{\Sigma}_{11} - \Delta_1)^{-1} - \hat{\Psi}_{21}(\hat{\Psi}_{11} - \Delta_1)^{-1}$, and the last q components of U are given by $\hat{\mu}_2 - \hat{\Sigma}_{21}(\hat{\Sigma}_{11} - \Delta_1)^{-1}\hat{\mu}_1 - \hat{\nu}_2 + \hat{\Psi}_{21}(\hat{\Psi}_{11} - \Delta_1)^{-1}\nu_1$. For H_2, U is obtained from a triangular portion of $\hat{\Sigma}_{22} - \hat{\Sigma}_{21}(\hat{\Sigma}_{11} - \Delta_1)^{-1}\hat{\Sigma}_{12} - \hat{\Psi}_{22} + \hat{\Psi}_{21}(\hat{\Psi}_{11} - \Delta_1)^{-1}\hat{\Psi}_{12}$, written as a vector. The reader is referred to Stocking and Lord [1973] for a further description of the computing procedure.

In order that U denote what is intended, i.e., an approximately normal random vector with mean $g(\mu,\nu,\Sigma,\Psi)$, both $\hat{\Sigma}_{11} - \Delta_1$ and $\hat{\Psi}_{11} - \Delta_1$ ought to be positive definite, yet theoretically this can fail. If the data turn out to be such that either $\hat{\Sigma}_{11} - \Delta_1$ or $\hat{\Psi}_{11} - \Delta_1$ is near singularity or has negative roots, one may regard this as an indication that, due to insufficient sample size, the measurement error has swamped out the information in the data estimating $\Sigma_{11} - \Delta_1$ or $\Psi_{11} - \Delta_1$. Evidence is presented in Stroud [1968, pp. 80-81] that with standardized achievement test data with $p = q = 1$ and sample sizes greater than 40, $\Sigma - \Delta$ (and hence $\Sigma_{11} - \Delta_1$) is almost certain to be positive definite. For the multiple regression case, perhaps larger sample sizes than this would be necessary to ensure that the anomaly would not occur.

The necessary formulas for computing W without using the Lord algorithm are given in the appendix. In this case one has the additional work of substituting the estimates $\hat{\mu}$, $\hat{\nu}$, $\hat{\Sigma}$, and $\hat{\Psi}$ into rather complicated formulas for the asymptotic covariance matrices. The computation of W has been carried out on the data described in the next section both using the Lord algorithm and using the asymptotic covariance matrix formulas. The results agree precisely to at least the number of significant digits retained in Table 1.

4. Analysis of Some Achievement Test Battery Scores

In this section we report the results of the study, mentioned in the introduction, based on grade 9 ITED (X_1 and Y_1) and grade 11 TAP (X_2 and Y_2) subtest scores. The comparison was between one specified school in Portland, Oregon and the group consisting of the other 11 high schools in that city.

A version of this study in which X_1, X_2, Y_1, and Y_2 were univariate (composite scores) has already been reported [Stroud, 1972]. The portion of these results concerning the tests of hypotheses H_1 and H_2 are included here for purposes of comparison with the multivariate tests.

The TAP subtests used were Composition, Reading, and Mathematics, in that order, and the ITED subtests were Social Concepts, Correctness of Expression, Quantitative Thinking, Reading (Social Studies), and Reading (Natural Sciences).

First the procedures were applied using as X_1 and X_2 the standard scores of all the girls in the specified school (School 1) and for Y_1 and Y_2 the scores of the girls in the rest of the city (Schools 2-12) taken together. Secondly, the same procedure is repeated for the boys. In the third and final application, the girls in the twelve schools (X_1, X_2) are compared with the boys (Y_1, Y_2). Table 1 shows the value of W for comparing (i) conditional mean vectors (hypothesis H_1) and (ii) conditional covariance matrices (hypothesis H_2) for both the multivariate data just described and the univariate case utilizing composite scores (averages over the subtests). Beside the value of W is given the corresponding value of P, the inverse of the approximating chi-square cumulative distribution function (P = 1 - significance level attained). The sample sizes are shown in Table 2.

The standard errors of measurement for the ITED subtests were taken to be 3.32, 3.16, 3.46, 3.32, and 3.46, respectively. These values were derived from the administrator's manual [Science Research Associates, 1963], and are expressed in the appropriate scaling of the Portland standardization. Subtest measurement errors were assumed to be uncorrelated with each other.

Insert Table 1 about here

It is seen from the significance levels in Table 1 that the results of the univariate and multivariate applications do not completely correspond with each other. Notice that in the school-versus-school comparisons (boys and girls separately) the multivariate test for covariance matrices reveals greater significance than the corresponding univariate test, but the multivariate test for mean vectors shows less significance. In the boy-versus-girl comparison, however, the pattern is reversed.

Insert Table 2 about here

In trying to account for these phenomena, we may note from Table 2, where the six 8-dimensional mean vectors are tabulated, that subtest scores for School 1 are consistently lower than those of the 2-12 group. However, if we compare boys and girls, we find that boys do better in some subtests (notably the quantitative) whereas girls do better in others (e.g., composition). Thus the composite scores used in the $p = q = 1$ analysis are appropriate for comparing schools, but not for comparing sexes because the sex-related differences will tend to be reduced in the averaging of subtest scores. In the school-versus-school comparison of mean vectors, most of the meaningful variation has been recorded in the composite score analysis; so that for example a chi-square of 11.92 on 2 degrees of freedom is registered as being more significant than a chi-square of 29.94 on 18

degrees of freedom, even though the difference of the chi-squares (18.02) slightly exceeds the difference in the degrees of freedom.

In the boy-versus-girl comparison, on the other hand, the multivariate analysis of mean vectors gives a more impressive result than the univariate for the reason (mentioned above) that the simple average is far from being the best linear combination demonstrating differences between boys and girls.

The interpretation of the results regarding conditional variances and covariance matrices is more difficult. The main factor is probably that residual variances based on predicting a single variable by a single variable cannot be expected to resemble too closely a 3×3 residual covariance matrix based on five predictors. One would guess that, with the school-versus-school comparisons, there are discrepancies in the residual covariance matrices which are washed out when we look at just the residual variance of the composite score. Regarding the boy-versus-girl comparison, an examination of the data has revealed that the difference between residual variances in the univariate analysis exceeds (in relative terms as well as absolute) the difference in residual variances of any of the three subtest scores in the multivariate results. This may very well be related to the interaction between sex and subtest content, but the pattern appears too complicated to give a detailed account here.

In conclusion, it would appear that the univariate and multivariate analyses taken together are more informative than either one would be alone. Although the techniques used to study the above data have been derived from the theory of inference, the author has the distinct impression that an honest attempt to get as good a feel for the data as possible is more fruitful than the making of statistical decisions such as accepting

or rejecting hypotheses. This supports the view of Dempster [1969] that the data-analytic approach to multivariate problems is often more sensible than a "solution" based on inference.

APPENDIX

A1. A Simplifying Transformation of the Parameter Space

The statistic W defined by (3.1) for testing either H_1 or H_2 may be computed as described in Section 3. This process involves obtaining $\hat{\Xi}$ by numerical differentiation of the function $g(\mu,\nu,\Sigma,\Psi)$ at the points in the parameter space specified by the estimates $\hat{\mu}$, $\hat{\nu}$, $\hat{\Sigma}$, and $\hat{\Psi}$. We will now proceed to derive the analytic form of Ξ. This yields an alternative method of computing W, in addition to providing a base for possible study of the properties of W.

We assume in the following derivations that the pretest measurement error covariance matrix is nonsingular. This assumption is not essential, but it makes possible a transformation which simplifies some of the terms in the matrix expressions. The derivations of formulas for the more general case may be carried out with straightforward modifications.

The transformation is that of rescaling the pretest variables to unit error of measurement, and is carried out as follows. Assume Δ_1 is nonsingular; then define $\tilde{T}$, $\tilde{U}$, $\tilde{E}$, and $\tilde{F}$ as transformed values of T, U, E, and F, respectively, after premultiplication by the matrix $\Gamma \equiv \begin{bmatrix} \Delta_1^{-\frac{1}{2}} & 0 \\ 0 & I \end{bmatrix}$, e.g., $\tilde{T} = \Gamma T$. Let $\tilde{X} = \tilde{T} + \tilde{E}$ and $\tilde{Y} = \tilde{U} + \tilde{F}$; then $\tilde{X} = \Gamma X$ and $\tilde{Y} = \Gamma Y$. Then $\tilde{T} \sim \eta(\tilde{\mu},\tilde{\Sigma}^*)$ and $\tilde{U} \sim \eta(\tilde{\nu},\tilde{\Psi}^*)$, where $\mu_1 = \Delta_1^{\frac{1}{2}}\tilde{\mu}_1$, $\mu_2 = \tilde{\mu}_2$, $\Sigma^*_{21} = \tilde{\Sigma}^*_{21}\Delta_1^{\frac{1}{2}}$, $\Sigma^*_{11} = \Delta_1^{\frac{1}{2}}\tilde{\Sigma}^*_{11}\Delta_1^{\frac{1}{2}}$, and $\Sigma^*_{22} = \tilde{\Sigma}^*_{22}$, with similar identities for the ν and Ψ quantities. Clearly the hypotheses (2.1) and (2.2) are unchanged when μ, ν, Σ^*, and Ψ^* are replaced by $\tilde{\mu}$,

$\tilde{\nu}$, $\tilde{\Sigma}^*$, and $\tilde{\psi}^*$, respectively. Let us therefore work with the ~ -quantities, forgetting the old ones, and the ~ symbol need be retained no longer since it is unnecessary. However, note that now E and F each have the covariance matrix $\Gamma\Delta\Gamma = \begin{bmatrix} I & 0 \\ 0 & \Delta_2 \end{bmatrix}$. The mean vectors and covariance matrices of X and Y are (μ,Σ) and (ν,ψ) , respectively. The hypotheses now read

$$\text{(A1)} \qquad H_1: \begin{cases} \Sigma_{21}(\Sigma_{11} - I)^{-1} - \psi_{21}(\psi_{11} - I)^{-1} = 0 \\ \mu_2 - \Sigma_{21}(\Sigma_{11} - I)^{-1}\mu_1 - \nu_2 + \psi_{21}(\psi_{11} - I)^{-1}\nu_1 = 0 \end{cases}$$

$$\text{(A2)} \qquad H_2: \quad \Sigma_{22} - \Sigma_{21}(\Sigma_{11} - I)^{-1}\Sigma_{12} - \psi_{22} + \psi_{21}(\psi_{11} - I)^{-1}\psi_{12} = 0 \quad .$$

In practice one carries out the transformation simply by dividing each pretest variable at the outset by its standard error of measurement; then one tests the hypothesis defined by (A1) and (A2).

A2. General Results for Asymptotic Covariance Matrices

We now state some general results which apply to problems of testing hypotheses concerning samples from two normal distributions. Let the hypothesis to be tested be $g(\mu,\nu,\Sigma,\psi) = 0$ and let the test statistic be $W = U'\hat{\Xi}^{-1}U$, as defined in (3.1). Consider m observations of $X \sim \mathcal{N}(\mu,\Sigma)$ and n observations of $Y \sim \mathcal{N}(\nu,\psi)$, all independent. We are concerned with asymptotic results which hold when m and n increase such that $n/m \to \rho$ as $n \to \infty$, where $0 < \rho < \infty$. Recall that

Ξ is an asymptotic approximation to the covariance matrix of U, and hence is of order n^{-1} (under regularity conditions). We define $\Omega = n\Xi$, where Ω is a function of μ, ν, Σ, ψ which does not depend on n. Then $W = nU'\hat{\Omega}^{-1}U$, where $\hat{\Omega}$ is obtained from Ω by replacing μ, ν, Σ, and ψ by $\hat{\mu}$, $\hat{\nu}$, $\hat{\Sigma}$, and $\hat{\psi}$, respectively. It is shown in Stroud [1971] that if g is twice differentiable with matrix of first derivatives of full rank then the distribution of W is asymptotically central (noncentral) chi-square when $g(\mu,\nu,\Sigma,\psi)$ is zero (nonzero), and that the components of Ω are given by

$$\text{(A3)} \qquad \omega(\alpha,\beta) = \rho\alpha_\mu'\Sigma\beta_\mu + \alpha_\nu'\psi\beta_\nu + 2\rho \operatorname{tr} \alpha_\Sigma\Sigma\beta_\Sigma\Sigma + 2 \operatorname{tr} \alpha_\psi\psi\beta_\psi\psi \quad ,$$

where α is defined as $g_i(\mu,\nu,\Sigma,\psi)$ (the i-th component of g when written as a vector), β is defined as $g_j(\mu,\nu,\Sigma,\psi)$, and by $\omega(\alpha,\beta)$ is meant the (i,j)-th component of Ω. Subscripts in (A3) denote partial differentiation; e.g., β_μ is the $(p+q)$-vector of partial derivatives of g_j with respect to the components of μ, and α_Σ is the $(p+q)\times(p+q)$ symmetric matrix of partial derivatives of g_i with respect to the components of Σ. In the latter case, since Σ is symmetric, the off-diagonal components of α_Σ include a factor of $\frac{1}{2}$ [see Stroud, 1971, formula 4.6; and Aitken, 1953]; derivatives with respect to symmetric matrices are defined this way for simplicity of the resulting formulas.

The final two sections of this appendix are concerned with deriving formulas for the $\omega(\alpha,\beta)$ when g is given by H_1 or H_2. Since H_2 is simpler, it is treated first.

A3. Components of Ξ for Testing Equality of Residual Covariance Matrices

Define Π as the function of the parameters representing the left-side of (A2); the hypothesis is $\Pi = 0$. A typical component of the matrix $\Omega = n\Xi$ is given by (A3), which may be written as

$$\omega_{ij,k\ell} = 2\rho \text{ tr } \pi_{\Sigma}^{ij}\Sigma\pi_{\Sigma}^{k\ell}\Sigma + 2 \text{ tr } \pi_{\psi}^{ij}\psi\pi_{\psi}^{k\ell}\psi \tag{A4}$$

(note that μ and ν are not involved), where, for example, π_{Σ}^{ij} is the matrix of partial derivatives of the (i,j)-th component of the matrix Π with respect to the components of Σ.

The partial derivatives π_{Σ}^{ij} and π_{ψ}^{ij} are evaluated with the aid of matrix differentials and the associated formulas for products $d(AB) = (dA)B + A(dB)$ and for inverses $d(A^{-1}) = -A^{-1}(dA)A^{-1}$ [see, e.g., Deemer and Olkin, 1951, results 5A13, 5A15, and 5B3]. If $Y = AXB$ (where all capitals denote matrices), the formula

$$\frac{\partial y_{ij}}{\partial X} = A'E_{ij}B' \quad ,$$

where E_{ij} consists of a "1" in position (i,j) and zeros elsewhere [Dwyer & MacPhail, 1948] is used to evaluate the matrix derivative. In case X is symmetric, the formula becomes

$$\frac{\partial y_{ij}}{\partial X} = \tfrac{1}{2}(A'E_{ij}B' + BE_{ij}'A) \quad ,$$

where $\partial y_{ij}/\partial X$ is defined with a factor of $\frac{1}{2}$ in the off-diagonal elements.

The differential of Π is now obtained. From (A2) it follows that

$$(A5)\quad d\Pi = d\Sigma_{22}-d\Sigma_{21}(\Sigma_{11}-I)^{-1}\Sigma_{12}-\Sigma_{21}(\Sigma_{11}-I)^{-1}d\Sigma_{12}+\Sigma_{21}(\Sigma_{11}-I)^{-1}d\Sigma_{11}(\Sigma_{11}-I)^{-1}\Sigma_{12}$$
$$- d\Psi_{22}+d\Psi_{21}(\Psi_{11}-I)^{-1}\Psi_{12}+\Psi_{21}(\Psi_{11}-I)^{-1}d\Psi_{12}-\Psi_{21}(\Psi_{11}-I)^{-1}d\Psi_{11}(\Psi_{11}-I)^{-1}\Psi_{12} \ .$$

The following partial derivatives are obtained:

$$(A6)\quad \begin{cases} \dfrac{\partial\pi^{ij}}{\partial\Sigma_{22}} = \dfrac{\partial(d\pi^{ij})}{\partial(d\Sigma_{22})} = \tfrac{1}{2}(E_{ij}+E'_{ij}) = \tfrac{1}{2}(E_{ij}+E_{ji}) \\ \dfrac{\partial\pi^{ij}}{\partial\Psi_{22}} = \dfrac{\partial(d\pi^{ij})}{\partial(d\Psi_{22})} = -\tfrac{1}{2}(E_{ij}+E_{ji}) \\ \dfrac{\partial\pi^{ij}}{\partial\Sigma_{11}} = \tfrac{1}{2}[(\Sigma_{11}-I)^{-1}\Sigma_{12}E_{ij}\Sigma_{21}(\Sigma_{11}-I)^{-1} + (\Sigma_{11}-I)^{-1}\Sigma_{12}E'_{ij}\Sigma_{21}(\Sigma_{11}-I)^{-1}] \\ \qquad = \tfrac{1}{2}(\Sigma_{11}-I)^{-1}\Sigma_{12}(E_{ij}+E_{ji})\Sigma_{21}(\Sigma_{11}-I)^{-1} \\ \dfrac{\partial\pi^{ij}}{\partial\Psi_{11}} = -\tfrac{1}{2}(\Psi_{11}-I)^{-1}\Psi_{12}(E_{ij}+E_{ji})\Psi_{21}(\Psi_{11}-I)^{-1} \\ \dfrac{\partial\pi^{ij}}{\partial\Sigma_{21}} = -(E_{ij}+E_{ji})\Sigma_{21}(\Sigma_{11}-I)^{-1} \\ \dfrac{\partial\pi^{ij}}{\partial\Psi_{21}} = (E_{ij}+E_{ji})\Psi_{21}(\Psi_{11}-I)^{-1} \ . \end{cases}$$

Remembering that the matrix $\partial\pi^{ij}/\partial\Sigma$ is evaluated with a factor of $\frac{1}{2}$ applied to all off-diagonal components, we may write it, using (A6), as follows, where the $q \times q$ symmetric matrix F_{ij} is defined by $F_{ij} = \frac{1}{2}(E_{ij}+E_{ji})$:

$$(A7)\qquad \pi_{\Sigma}^{ij} = \begin{bmatrix} (\Sigma_{11}-I)^{-1}\Sigma_{12}F_{ij}\Sigma_{21}(\Sigma_{11}-I)^{-1} & -(\Sigma_{11}-I)^{-1}\Sigma_{12}F_{ij} \\ -F_{ij}\Sigma_{21}(\Sigma_{11}-I)^{-1} & F_{ij} \end{bmatrix} .$$

Since Σ may be written as

$$\Sigma = \begin{bmatrix} (\Sigma_{11}-I)+I & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} ,$$

it follows from (A7). by straightforward evaluation, that

$$(A8)\qquad \operatorname{tr} \pi_{\Sigma}^{ij}\Sigma\pi_{\Sigma}^{k\ell}\Sigma = \operatorname{tr} F_{ij}A^{(1)}F_{k\ell}A^{(1)}$$

$$= \tfrac{1}{2}(a_{ik}^{(1)}a_{j\ell}^{(1)} + a_{i\ell}^{(1)}a_{jk}^{(1)})$$

[cf. Stroud, 1971, formula 3.1], where

$$(A9)\qquad A^{(1)} = \Sigma_{22}-\Sigma_{21}(\Sigma_{11}-I)^{-1}\Sigma_{12} + \Sigma_{21}(\Sigma_{11} - I)^{-2}\Sigma_{12} ,$$

and $a_{ik}^{(1)}$ stands for the (i,k) -th component of $A^{(1)}$.

Define $A^{(2)}$ in terms of ψ as in (A9). If we substitute into (A4) the formula (A8) and the analogous formula involving ψ and $A^{(2)}$, the following result is yielded:

$$\omega_{ij,k\ell} = \frac{n}{m}\left(a_{ik}^{(1)}a_{j\ell}^{(1)} + a_{i\ell}^{(1)}a_{jk}^{(1)}\right) + \left(a_{ik}^{(2)}a_{j\ell}^{(2)} + a_{i\ell}^{(2)}a_{jk}^{(2)}\right)$$

$$(1 \le i \le j \le q, \quad 1 \le k \le \ell \le q) \quad ,$$

where ρ is taken here as equal to n/m .

In the case $q = 1$, Ω equals the scalar $\omega_{11,11}$. $U = \hat{\Pi}$ is also a scalar, so it is easy to write down the following formula for the statistic W :

$$W = \frac{\frac{1}{2}[\hat{\Sigma}_{22}-\hat{\Sigma}_{21}(\hat{\Sigma}_{11}-I)^{-1}\hat{\Sigma}_{12}-\hat{\Psi}_{22}+\hat{\Psi}_{21}(\hat{\Psi}_{11}-I)^{-1}\hat{\Psi}_{12}]^2}{m^{-1}[\hat{\Sigma}_{22}-\hat{\Sigma}_{21}(\hat{\Sigma}_{11}-I)^{-1}\hat{\Sigma}_{12}+\hat{\Sigma}_{21}(\hat{\Sigma}_{11}-I)^{-2}\hat{\Sigma}_{12}]^2+n^{-1}[\hat{\Psi}_{22}-\hat{\Psi}_{21}(\hat{\Psi}_{11}-I)^{-1}\hat{\Psi}_{12}+\hat{\Psi}_{21}(\hat{\Psi}_{11}-I)^{-2}\hat{\Psi}_{12}]^2} \quad .$$

A4. Testing for Equality of Regressions

The development of this section parallels that of the preceding section, but the presentation is somewhat more condensed to save space. This time the quantity $g(\mu,\nu,\Sigma,\Psi)$ is the vector of dimension $pq + q$ whose components are the components of the $q \times p$ matrix

$$\Phi = \Sigma_{21}(\Sigma_{11}-I)^{-1} - \Psi_{21}(\Psi_{11}-I)^{-1}$$

and the $q \times 1$ vector

$$\Lambda = \mu_2-\Sigma_{21}(\Sigma_{11}-I)^{-1}\mu_1-\nu_2+\Psi_{21}(\Psi_{11}-I)^{-1}\nu_1 \quad .$$

It is straightforward to show that the matrix of partial derivatives of the transformation from (μ,ν,Σ,ψ) to (ϕ,Λ) has full rank of $pq + q$, by showing that the Jacobian of the transformation from (Σ_{21},μ_2) to (ϕ,Λ) is nonzero [see Deemer & Olkin, 1951, Theorem 3.5].

The $pq + q$ components of U are the components ϕ^{ij} $(1 \le i \le q,\ 1 \le j \le p)$ of the matrix ϕ and the components λ^{i} $(1 \le i \le q)$ of the vector Λ when (μ,ν,Σ,ψ) is replaced by its estimator $(\hat{\mu},\hat{\nu},\hat{\Sigma},\hat{\psi})$.

Each component of Ω, as given by (A3), is one of three types, according to whether α and β are:

(i) both components of ϕ

(ii) both components of Λ

(iii) one a component of ϕ and the other of Λ.

Let the components of Ω of these types be denoted, respectively, by $\omega_{ij,k\ell}$ $(i,k = 1,\ldots,q;\ j,\ell = 1,\ldots,p)$, $\omega_{i,k}$ $(i,k = 1,\ldots,q)$, and $\omega_{i,k\ell}$ $(i,k = 1,\ldots,q;\ \ell = 1,\ldots,p)$. The general formula (A3) may be rewritten as follows:

$$\text{(A10)}\quad \begin{cases} \omega_{ij,k\ell} = \rho(\phi_{\mu}^{ij})'\Sigma\phi_{\mu}^{k\ell} + (\phi_{\nu}^{ij})'\psi\phi_{\nu}^{k\ell} + 2\rho\ \mathrm{tr}\ \phi_{\Sigma}^{ij}\Sigma\phi_{\Sigma}^{k\ell}\Sigma + 2\ \mathrm{tr}\ \phi_{\psi}^{ij}\psi\phi_{\psi}^{k\ell}\psi\ , \\ \omega_{i,k} = \rho(\lambda_{\mu}^{i})'\Sigma\lambda_{\mu}^{k} + (\lambda_{\nu}^{i})'\psi\,\lambda_{\nu}^{k} + 2\rho\ \mathrm{tr}\ \lambda_{\Sigma}^{i}\Sigma\lambda_{\Sigma}^{k}\Sigma + 2\ \mathrm{tr}\ \lambda_{\psi}^{i}\psi\lambda_{\psi}^{k}\psi\ , \\ \omega_{i,k\ell} = \rho(\lambda_{\mu}^{i})'\Sigma\phi_{\mu}^{k\ell} + (\lambda_{\nu}^{i})'\psi\phi_{\nu}^{k\ell} + 2\rho\ \mathrm{tr}\ \lambda_{\Sigma}^{i}\Sigma\phi_{\Sigma}^{k\ell}\Sigma + 2\ \mathrm{tr}\ \lambda_{\psi}^{i}\psi\phi_{\psi}^{k\ell}\psi\ , \end{cases}$$

where $i,k = 1,\ldots,q$; $j,\ell = 1,\ldots,p$, and $\rho = n/m$. The next step is to obtain formulas for expressions like ϕ_{Σ}^{ij}, λ_{μ}^{i}, and λ_{Σ}^{i} and then find simplified expressions for the terms of the right sides of (A10).

To get ϕ_Σ^{ij}, first write the differential of Φ :

$$d\Phi = d\Sigma_{21}(\Sigma_{11}-I)^{-1}-\Sigma_{21}(\Sigma_{11}-I)^{-1}d\Sigma_{11}(\Sigma_{11}-I)^{-1}-d\Psi_{21}(\Psi_{11}-I)^{-1}$$

$$+ \Psi_{21}(\Psi_{11}-I)^{-1}d\Psi_{11}(\Psi_{11}-I)^{-1} \quad .$$

Hence, after differentiation with respect to submatrices, incorporating the factor of $\frac{1}{2}$ where necessary, one gets

(A11)
$$\phi_\Sigma^{ij} = \begin{bmatrix} -\frac{1}{2}(\Sigma_{11}-I)^{-1}(\Sigma_{12}E_{ij}+E'_{ij}\Sigma_{21})(\Sigma_{11}-I)^{-1} & \frac{1}{2}(\Sigma_{11}-I)^{-1}E'_{ij} \\ \frac{1}{2}E_{ij}(\Sigma_{11}-I)^{-1} & 0 \end{bmatrix} .$$

By straightforward calculation and noting $(\Sigma_{11}-I)^{-1}\Sigma_{11}(\Sigma_{11}-I)^{-1} = (\Sigma_{11}-I)^{-1} + (\Sigma_{11}-I)^{-2}$, we obtain from (A11) that

$$\operatorname{tr} \phi_\Sigma^{ij}\Sigma\phi_\Sigma^{k\ell}\Sigma = \tfrac{1}{2} \operatorname{tr} E_{ij}C^{(1)}E'_{k\ell}A^{(1)} + \tfrac{1}{2} \operatorname{tr} E_{ij}B^{(1)}E_{k\ell}B^{(1)} \quad ,$$

where, for $\alpha = 1,2$, the matrices $A^{(\alpha)}$ are defined by (A9), and the following notation is introduced:

$$B^{(1)} = (\Sigma_{11}-I)^{-2}\Sigma_{12} \quad , \quad C^{(1)} = (\Sigma_{11}-I)^{-1}+(\Sigma_{11}-I)^{-2} \quad ,$$

with $B^{(2)}$ and $C^{(2)}$ defined analogously for Ψ. Hence

$$\operatorname{tr} \phi_\Sigma^{ij}\Sigma\phi_\Sigma^{k\ell}\Sigma = \tfrac{1}{2}(c_{j\ell}^{(1)}a_{ki}^{(1)} + b_{jk}^{(1)}b_{i\ell}^{(1)}) \quad ,$$

and a similar formula holds involving Ψ.

Since Φ does not involve μ or ν, the first two terms in the formula for $\omega_{ij,k\ell}$ vanish, and hence

$$\omega_{ij,k\ell} = \rho(c^{(1)}_{j\ell}a^{(1)}_{ki} + b^{(1)}_{jk}b^{(1)}_{\ell i}) + (c^{(2)}_{j\ell}a^{(2)}_{ki} + b^{(2)}_{jk}b^{(2)}_{\ell i})$$

$$(i,k = 1,\ldots,q;\ j,\ell = 1,\ldots,p) \quad .$$

The following result is obtained for the differential $d\Lambda$:

$$d\Lambda = d\mu_2 - \Sigma_{21}(\Sigma_{11}-I)^{-1}d\mu_1 - d\Sigma_{21}(\Sigma_{11}-I)^{-1}\mu_1 + \Sigma_{21}(\Sigma_{11}-I)^{-1}d\Sigma_{11}(\Sigma_{11}-I)^{-1}\mu_1$$

$$- d\nu_2 + \Psi_{21}(\Psi_{11}-I)^{-1}d\nu_1 + d\Psi_{21}(\Psi_{11}-I)^{-1}\nu_1 - \Psi_{21}(\Psi_{11}-I)^{-1}d\Psi_{11}(\Psi_{11}-I)^{-1}\nu_1 \quad .$$

Using the fact that when E_{ij} contains only one column it may be written as e_i, defined as a column vector with "1" in the i-th position and zeros elsewhere, one may obtain

$$\text{(A12)} \quad \lambda^i_\Sigma = \begin{bmatrix} \frac{1}{2}[(\Sigma_{11}-I)^{-1}(\Sigma_{12}e_i\mu_1' + \mu_1 e_i'\Sigma_{21})(\Sigma_{11}-I)^{-1}] & -\frac{1}{2}(\Sigma_{11}-I)^{-1}\mu_1 e_i' \\ -\frac{1}{2}e_i\mu_1'(\Sigma_{11}-I)^{-1} & 0 \end{bmatrix}$$

and

$$\text{(A13)} \quad \lambda^i_\mu = \begin{bmatrix} -(\Sigma_{11}-I)^{-1}\Sigma_{12}e_i \\ e_i \end{bmatrix} \quad .$$

It may be noted that formulas (A11) and (A12) are identical, except that in (A12) the matrix $(-e_i\mu_1')$ replaces E_{ij} of (A11). Thus (A12) implies

$$(A14)\quad \operatorname{tr} \lambda_{\Sigma}^{i}\Sigma\lambda_{\Sigma}^{k}\Sigma = \tfrac{1}{2} \operatorname{tr} e_i\mu_1'C^{(1)}{}_{\mu_1}e_k'A^{(1)} + \tfrac{1}{2} \operatorname{tr} e_i\mu_1'B^{(1)}e_k\mu_1'B^{(1)}$$

$$= \tfrac{1}{2}(\mu_1'C^{(1)}{}_{\mu_1})a_{ik}^{(1)} + \tfrac{1}{2}(\mu_1'B^{(1)})_i(\mu_1'B^{(1)})_k \quad .$$

By further calculation it may be seen that relation (A13) implies

$$(A15)\qquad (\lambda_{\mu}^{i})'\Sigma\lambda_{\mu}^{k} = e_i'A^{(1)}e_k = a_{ik}^{(1)} \quad .$$

Substitution of (A14), (A15) and analogous expressions involving Ψ and ν into (A10) yields the following formula:

$$\omega_{i,k} = \rho[(\mu_1'C^{(1)}{}_{\mu_1}+1)a_{ik}^{(1)} + (\mu_1'B^{(1)})_i\,(\mu_1'B^{(1)})_k]$$

$$+ [(\nu_1'C^{(2)}{}_{\nu_1}+1)a_{ik}^{(2)} + (\nu_1'B^{(2)})_i(\nu_1'B^{(2)})_k]$$

$$(i,k = 1,\ldots,q) \quad .$$

To get the formula for $\omega_{i,k\ell}$, it is straightforward to obtain

$$\operatorname{tr} \lambda_{\Sigma}^{i}\Sigma\phi_{\Sigma}^{k\ell}\Sigma = -\tfrac{1}{2} \operatorname{tr} e_i\mu_1'C^{(1)}E_{k\ell}'A^{(1)} - \tfrac{1}{2} \operatorname{tr} e_i\mu_1'B^{(1)}E_{k\ell}B^{(1)}$$

$$= -\tfrac{1}{2}\, a_{ki}^{(1)}(\mu_1'C^{(1)})_{\ell} - \tfrac{1}{2}\, b_{i\ell}^{(1)}(\mu_1'B^{(1)})_k \quad .$$

Then, since $\phi_{\mu}^{k\ell} = \phi_{\nu}^{k\ell} = 0$, one obtains

$$\omega_{i,k\ell} = -\rho[a_{ki}^{(1)}(\mu_1'C^{(1)})_{\ell} + b_{i\ell}^{(1)}(\mu_1'B^{(1)})_k]$$

$$- [a_{ki}^{(2)}(\nu_1'C^{(2)})_{\ell} + b_{i\ell}^{(2)}(\nu_1'B^{(2)})_k]$$

$$(i,k = 1,\ldots,q;\ \ell = 1,\ldots,p) \quad .$$

Expressions for all the components of $\Omega = n\Xi$ have now been displayed.

When X_2 and Y_2 are univariate ($q = 1$, i.e., multiple but not multivariate regression), one can present a reasonable looking formula for Ω as a bordered matrix, rather than simply a collection of separate formulas for components such as appears above. Eliminate superscripts and write $a_\alpha \equiv A^{(\alpha)} = a_{11}^{(\alpha)}$, $B_\alpha \equiv B^{(\alpha)}$, and $C_\alpha \equiv C^{(\alpha)}$ for $\alpha = 1,2$. The a_α are scalars, the B_α are $p \times 1$ column vectors and the C_α are $p \times p$ symmetric matrices. Then define

$$G_\alpha = a_\alpha C_\alpha + B_\alpha B_\alpha' \qquad (\alpha = 1,2) \quad .$$

It is then straightforward to write the formula for Ω ,

$$\Omega = \frac{1}{m}\begin{bmatrix} nG_1+mG_2 & -nG_1\mu_1-mG_2\nu_1 \\ -n\mu_1'G_1-m\nu_1'G_2 & n(a_1+\mu_1'G_1\mu_1)+m(a_2+\nu_1'G_2\nu_1) \end{bmatrix} \quad .$$

If $p = 1$ as well (simple regression), the formula for Ω^{-1} may be easily written down. The corresponding formula for $W = nU'\hat{\Omega}^{-1}U$ may be found in Stroud [1972].

References

Aitken, A. C. A note on trace-differentiation and the Ω -operator. Proceedings of the Edinburgh Mathematical Society (2), 1953, 10, 1-4.

Anderson, T. W. An introduction to multivariate statistical analysis. New York: Wiley, 1958.

Cochran, W. G. Errors of measurement in statistics. Technometrics, 1968, 10, 637-666.

Cronbach, L. G. & Gleser, G. C. Psychological tests and personnel decisions. (2nd ed.) Urbana, Ill.: University of Illinois Press, 1965.

Deemer, W. L. & Olkin, I. The Jacobians of certain matrix transformations useful in multivariate analysis, based on lectures by P. L. Hsu. Biometrika, 1951, 38, 345-367.

Dempster, A. P. Elements of continuous multivariate analysis. Reading, Mass.: Addison-Wesley, 1969.

Dwyer, P. S. & MacPhail, M. S. Symbolic matrix derivatives. Annals of Mathematical Statistics, 1948, 19, 517-534.

Johnson, P. O. & Jackson, R. W. B. Modern statistical methods: descriptive and inductive. Chicago: Rand McNally, 1959.

Kendall, M. G. Regression, structure, and functional relationships, Part I. Biometrika, 1951, 38, 11-25.

Lord, F. M. Large-sample covariance analysis when the control variable is fallible. Journal of the American Statistical Association, 1960, 55, 307-321.

Lord, F. M. Automated hypothesis tests and standard errors for nonstandard problems. Research Bulletin 72-42. Princeton, N.J.: Educational Testing Service, 1972.

Lord, F. M. & Novick, M. R. Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley, 1968.

Madansky, A. The fitting of straight lines when both variables are subject to error. Journal of the American Statistical Association, 1959, 54, 173-205.

Moran, P. A. P. Estimating structural and functional relationships. Journal of Multivariate Analysis, 1971, 1, 232-255.

Science Research Associates. The Iowa tests of educational development: manual for the school administrator. Chicago: Science Research Associates, Inc., 1963.

Stocking, M. & Lord, F. M. AUTEST - Program to perform automated hypothesis tests for nonstandard problems. Research Memorandum RM-73-00. Princeton, N.J.: Educational Testing Service, 1973, in press.

Stroud, T. W. F. Comparing conditional distributions under measurement errors of known variances. Unpublished Ph.D. dissertation, Stanford University, 1968.

Stroud, T. W. F. On obtaining large-sample tests from asymptotically normal estimators. Annals of Mathematical Statistics, 1971, 42, 1412-1424.

Stroud, T. W. F. Comparing conditional means and variances in a regression model with measurement errors of known variances. Journal of the American Statistical Association, 1972, 67, 407-412.

Wald, A. Tests of statistical hypotheses concerning several parameters when the number of observations is large. Transactions of the American Mathematical Society, 1943, 54, 426-482.

TABLE 1

Values of the Test Statistic W and Corresponding Probability Point P Based on Analysis of Composite Scores (p = q = 1) and Vector Scores (p = 5, q = 3)

	Comparing Conditional Covariance Matrices				Comparing Conditional Mean Vectors			
	Composite Scores		Vector Scores		Composite Scores		Vector Scores	
	W	P	W	P	W	P	W	P
Schools 1 versus 2-12 (Girls)	1.04	.77	13.01	.957	11.92	.9974	29.94	.962
Schools 1 versus 2-12 (Boys)	1.00	.68	12.88	.955	3.44	.82	20.57	.70
Girls versus Boys	6.89	.9913	10.37	.890	19.06	.99992	633.27	>.99999
Degrees of Freedom for W	1		6		2		18	

TABLE 2

Sample Sizes and Mean Subtest Scores

	Girls 1	Girls 2-12	Boys 1	Boys 2-12	All Girls	All Boys
Sample Size	195	1779	167	1754	1974	1921
Grade 9 ITED:						
Soc. Concepts	47.52	49.56	49.98	51.51	49.36	51.38
Corr. Expr.	50.32	53.12	46.75	48.67	52.84	48.51
Quantitative	48.14	49.04	50.29	51.20	48.96	51.12
Reading (SS)	49.73	51.07	49.53	50.77	50.93	50.66
Reading (NS)	48.53	49.81	48.73	50.21	49.69	50.08
Grade 11 TAP:						
Composition	49.57	52.51	44.78	47.13	52.22	46.92
Reading	47.53	50.03	47.98	49.35	49.78	49.23
Mathematics	46.69	49.09	50.96	53.24	48.85	53.04

Acknowledgments

The main results in this paper were first obtained as part of the author's Ph.D. dissertation at Stanford University, under the supervision of Professor Ingram Olkin and with the partial support of National Science Foundation grants GP-6681 and GP-27210. The author wishes to thank the Portland School District Administration for the use of their data. The author is grateful for the comments of Dr. Frederic M. Lord.